DOCUMENT RESUME

ED 081 487 PS 006 742

AUTHOR Wolf, Robert L.
TITLE A New Look at Evaluating Early Childhood Education Programs.
INSTITUTION Illinois Univ., Urbana. Center for Instructional Research and Curriculum Evaluation.
PUB DATE 73
NOTE 14p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Early Childhood Education; Educational Accountability; Educational Administration; *Educational Environment; *Formative Evaluation; Psychometrics; Reliability; Student Teacher Relationship; *Summative Evaluation; *Teacher Evaluation; Validity

ABSTRACT

It is argued that early childhood education programs should be evaluated in terms of criteria other than measurement and testing procedures. Programs should be judged on criteria such as responsiveness to clients--the children, parents, and communities they serve. Teachers trained in observational and evaluative techniques should conduct ongoing assessment of children. The importance of teacher self-evaluation to examine effectiveness is discussed in relation to the emotional climate of the program as established by administrators. Summative evaluation, which assesses the overall worth of a program at an end point, is distinguished from formative evaluation, which focuses on strengths and weaknesses of the program and leads to strategies for improvement. The latter type of assessment is recommended for education, using feedback from parents, other teachers, administrators, paraprofessionals, the children, and the community. The implementation of formative evaluation procedures would affect the nature of accountability for program success or failure. (DP)

ED 081487

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A NEW LOOK AT EVALUATING EARLY CHILDHOOD EDUCATION PROGRAMS

by

Robert L. Wolf
Center for Instructional Research
and Curriculum Evaluation
University of Illinois at
Urbana-Champaign
May, 1973

The evaluation of early childhood education programs has for a long time focused on the outcomes--the products of learning--as measured by the performance of the children involved. Now we must begin to judge early childhood education programs on other criteria such as responsiveness to clients--the children, parents, and communities they serve. The criteria for judgment need not be new, but a new emphasis should be attached to them. If early childhood education programs are to become more accountable to children, they ought to be judged by what they offer children. Goals which have become implicit--and dormant--become salient.

What we expect of children is contingent upon what we provide for them. It is our performance which should now be assessed. And what we need to provide is an enriched educational process. In evaluating programs for young children, we should consider more than the children themselves (although this is still important) and begin to judge the people who work and teach in these programs. In other words, evaluating early childhood education programs involves more than evaluating the performance of children; it involves the creation of an attitude in those who work with young children which will make them want more evaluative information regarding their own effectiveness. This broader perspective of evaluation then implies not only developing better

PS 006742

means to study and judge children, but the self-evaluation of those who staff early childhood programs.

Measurement vs. Evaluation

Historical Background

For a long time in education evaluation has been synonymous with measurement and testing procedures. To some extent this is still true today. But as the nature of early childhood education begins to change, evaluation should become more broadly responsive. The utilization of testing procedures as the sole criterion for judging the worth of programs needs to be challenged for several reasons.

Problems of Testing

For one thing, employing tests for young children under the ages of 7 or 8 presents serious problems related to the validity and reliability in the use and meaning attached to the results of such testing procedures. In addition to the more obvious conditions of lack of concentration, interest, and general awareness of why the tests are given, young children do not take tests seriously. This casual attitude violates one of the most basic assumptions underlying educational testing--that is, the child is maximally motivated to perform to the best of his ability. When looking at the well documented fluctuation of IQ scores over chronological age during the first six years of life one becomes cognizant of the unreliability of such measures. A possible explanation for this variance in scores (the fluctuations occur in both directions over different testing intervals) is that distinct tasks

are assigned to measure the construct of intelligence. Working with blocks is not the same as completing a paper and pencil test and so the operational definition of intelligence really changes over both situations. We often fail to acknowledge this change, however, and we rarely discuss the consequences of it.

The problem of validity in regard to testing young children is best exemplified in achievement testing. Both achievement and intelligence testing are often employed in programs for young children. Achievement tests are most often used to predict future performance. In the case of young children this sole use is extremely invalid.

Qualities not Easily Tested

In addition to the problems of validity and reliability when using tests with young children, there is the more profound dilemma of not being able to measure some of the more desirable qualities that early childhood programs strive for. There are currently no tests to appropriately measure qualities such as gentleness, courage, caring, commitment, sense of humor, thoughtfulness, sense of wonder, inquiry, generosity, etc., and yet these attitudes and behaviors have special significance for programs and people who concern themselves with early childhood education.

Testing and Oversimplification

To judge the success or failure of programs for young children on the basis of performance on tests, does not do justice to those programs--and all who are involved in or affected by them, nor does it lead to a meaningful understanding of those programs, which includes identifying strengths and weaknesses. Test results certainly tell us very little about how programs can be

improved. For example, let us for the moment look at the evaluation of Head Start.

Head Start was created to be a preschool compensatory education program. It was emphasized general child development and not the teaching of skills, per se. Most Head Start programs have been permissive-enrichment programs, characterized by their whole-child-orientation, their strategy of watching and waiting, and the resultant low degree of structure.

The Westinghouse evaluation study of Head Start picked at random some 104 centers (out of more than 1,200 centers throughout the nation) and all children eligible to enter each center were identified. Children who attended Head Start were chosen and placed in groups of eight. A carefully matched comparison group of children who did not attend Head Start was also selected. The children in both groups were extensively tested during the 1968-69 school year. Since the program began in the 1965 school year, it was possible to compare program and non-program in grades one, two, and three.

The evaluative study found that there were small but significant differences in favor of full-year Head Start children at the beginning of grade one on the Metropolitan Readiness Test (a generalized measure of learning readiness). But Head Start children at the beginning of grade two did not score significantly higher than the controls on the Stanford Achievement Test.

Through an almost exclusive reliance on the statistical treatment of standardized test scores and other outcome measures, the complexity of Head Start was reduced and simplified so that a decision as to the worth of those programs could be made. Parents, objected to the evaluations by stating that the tests used gave them little understanding about the programs,

and at the same time were culturally biased against their children. Even policy makers were dissatisfied on the ground that the information in the evaluation report was incomplete. It is clear that all relevant information leading to a true understanding of Head Start was not sought. Therefore, program decisions based on these evaluation reports are at best ill informed, and at worst irrelevant.

A child's school environment consists of many elements ranging from the desk he sits at to the child who sits next to him, and includes the teacher as well as the materials. A statistical treatment of test scores can give only fragmentary evidence of this environment. Great collections of numbers (totals, averages, percentages) such as those found in the evaluation of Head Start blur and obscure rather than sharpen and illuminate the range of variation they hope to represent.

Beyond Tests

The essential argument, as exemplified above, is that test scores provide little understanding of the complexity involved in early childhood education programs. The tests tell us nothing of the environment created by those who staff such programs. Tests also tell us little of the effectiveness, commitment, and skill of the teachers and workers in programs for young children. Worst of all, tests provide us with meager understanding of the growth and development of young children. It seems that we need to destroy the myth of testing, and in the process move away from measuring children and start developing meaningful understandings of them.

ED 060742

Some Other Inquiries Into Evaluation

Formative vs. Summative Evaluation

Before one can begin to discuss understanding as a goal of evaluation, one must first define two distinct functions that evaluation generally serves. These distinctions arise over the issues of formative and summative evaluation. Summative evaluation examines the overall worth of the program. It usually occurs when a program is nearing its completion. The results of such an evaluation frequently arrive after the program is terminated, therefore the evaluative information does not help the people in that program alter the program so that it may become more effective. Formative evaluation, on the other hand, is a continuous process. It would involve all members of an early childhood program as they systematically gathered information about themselves and the children in their charge. Formative evaluation implies identifying program strengths and weaknesses, and developing strategies to build upon those strengths and ameliorate those weaknesses. It is a dynamic evaluation effort that takes time, energy, and commitment, but inevitably leads to a greater understanding of the program's complexity. And this understanding becomes vital in making the program as effective as it can be for young children. If an evaluation effort does not help the staff and children improve upon what they are doing, then it is not worth the time, cost, and energy expended.

Within this formative evaluation context let us examine two crucial aspects of an evaluation effort--teacher self-evaluation, and more sensible ways to evaluate children. It is my belief, that because of the importance of their task, teachers of young children have a moral obligation to continuously question their own effectiveness so that they can provide the most meaningful environment and create the most sensible educational opportunities

that will nourish the growth (physical, mental, and social) and development of the young children in their programs. Similarly, those teachers need to understand their children so that the environment and opportunities they provide can best meet the diverse individual needs which exist.

Teacher Self-Evaluation

Teachers in early childhood education programs need to examine their effectiveness in providing valuable experiences for young children. In examining such effectiveness teachers will need to rely on the perceptions of children, parents, other teachers, and supervisors in contact with the program in question. These perceptions will give the teacher a better understanding of how other people view both the program and her effectiveness in that program. These various data sources represent the audiences that are served by or involved in early childhood programs.

Young children are refreshingly candid. When asked, they will provide a teacher with honest feelings about their classroom experience. Children will let a teacher know what they like and what they don't like--they can even explain why. Such information becomes crucial to teachers who wish to meet the individual needs of the children she is responsible for. Parents too can provide a teacher with meaningful information. The wise teacher will want to know what parents are feeling. She will want to learn how to become more responsive to these parental feelings, and help meet the needs of parents as well as children. Finally, other teachers and supervisors (principal, Head Start director, etc.), can provide professional criticism of a teacher's effectiveness in a program. By observing in a class, or interviewing children, other staff members can provide a teacher with suggestions and strategies for

improving practice.

There are at least two ways by which such suggestions and strategies can evolve. The teacher whose classroom program is being observed can provide the observer with a list of goals and objectives that are being strived for. Observers can then provide information as to whether the desired intentions of the self-evaluating teacher are being realized. This will enable the teacher to know what goals are being achieved and which are not, and also may provide other descriptive or explanatory information. The second method of observation is somewhat distinct. The teacher whose classroom is being observed provides no information as to what should be occurring. The observer then visits the classroom without a preconceived notion of what to look for. The evaluation becomes goal free, and simply describes what is going on according to the observer. The teacher can use this information to determine whether what is perceived as occurring in her classroom is what she really wishes to be occurring. If not, modifications can be sought.

To reiterate, it is important for any teacher of young children who wishes to become more responsive to the needs of the individual to receive evaluative feedback on her progress and effectiveness. In addition to developing introspective and reflective skills such teachers need to encourage their children, colleagues, parents, supervisors and other support staff to provide them with evaluative data as to their own strengths and weaknesses. The perceptions of these other individuals may give teachers greater insight into ways to improve practice and, as a result, become more effective. Self-evaluation can be seen as an effort to grow. Its purpose is not to assign blame, but to allow a teacher to learn more about herself

and her practice, so that she can better achieve the goals which are important to her. But evaluation is threatening. It is difficult to learn of one's shortcomings without engaging in defensive behavior. I have found that teachers' feelings toward evaluation are affected by the atmosphere of the school or program they work in. In environments where teachers feel more supported their attitudes toward evaluation are more open and they are more willing to engage in self-evaluation activities (the converse of this is also true). Teachers need to constructively channel feedback they receive so that it leads to improved practice--for this they need encouragement and support. Program administrators play a crucial role in creating an atmosphere that is conducive to self-evaluation efforts.

Building principals or program directors should try to give teachers more time for implementing new evaluation procedures, since lack of time seems to be a common problem for teachers. Similarly, administrators can begin to evaluate their own effectiveness and secure the support of their teachers and other staff in this effort, thereby providing a self-evaluation model for teachers to emulate.

Teachers should view themselves as an important audience for any results of classroom evaluation, such as judgments of children about their experiences, and feedback from parents and community groups. This evaluative information may be processed alone, but the administrator ought to encourage open interchange among the staff.

Thus, in the final analysis, it is the program administrator who has the best opportunity to create a climate that nurtures evaluation efforts. If this atmosphere is open and unthreatening, teacher self-evaluation is likely to occur with increasing frequency. Problems need to be discussed

and worked through, not harbored or suppressed. Administrators need to give teachers more autonomy so that teachers can feel more responsible to themselves, as well as others. Teachers need to be rewarded for their efforts in evaluating their own programs and administrators can have an important role in setting up such a reward system. The reward for the administrator will be worth his effort: he will have better information about what is happening in his program, and a staff that is monitoring its own efforts. A program in early childhood education where evaluation is an important part of everyone's concern is likely to be responsive to the needs of the young children it serves.

Toward Understanding Children

Being more responsive to young children through evaluation involves more than teachers willing to engage in self-evaluation efforts, it involves developing more sensible ways to evaluate the children themselves. As stated throughout this article early childhood educators need to develop more meaningful understandings of the children who occupy places in their programs. It has been argued here that tests and other measurement devices provide only limited information on the growth and development of young children, and at times, even this limited information tends to be both invalid and unreliable.

Teachers need to develop understanding of young children by developing more meaningful relationships with them. Out of such relationships evolves a dialogue between teacher and child that provides insight into the way the child is growing, learning, and developing as a person. Teachers need to be more skillful in observing, describing and inferring from the actions and

interactions of children what direction their growth is taking. It is in the process of early childhood education where indications of such development become apparent.

Most teachers can recognize when a child is engaged in an activity that he is enthusiastic about. Teachers can also perceive mindlessness. Teachers need to determine, with a series of questions, whether purposeful learning is occurring. It is important to note that while the teacher judges the nature of a child's experience, she must be cognizant of its educational value. Young children will often engage in some activity which is emotionally pleasing but that lacks educational purpose. It is up to the teacher to perceive this and try to provide some with a new challenge or intellectual stimulation.

It is impossible, however, for a teacher to be able to observe, intervene, and document all the important transactions for each child. The teacher needs to rely on the observations she can make and the descriptions which the children provide themselves. Children should be encouraged to talk about the activities they engage in; they should also be encouraged to keep a record of their work (putting all their work in one place over a period of time is the kind of record I am thinking about). Young children can be shown that this kind of record can help them in their own self-evaluation efforts. The strategy is inductive in that teachers should synthesize the fragments of information which result from their observations and interactions to understand how the child is developing and progressing.

To carry out this kind of evaluation strategy teachers will need to utilize a repertoire of skills, heretofore not required, honored, or rewarded.

For example, teachers of young children should know how to ask questions that are delicate and non-threatening, yet at the same time provacative. In addition to question asking skills, teachers also need to develop and use listening skills. Being able to listen to what a child is really saying when he talks becomes significant when the listener is attempting to acquire a meaningful understanding of him. In this respect, early childhood educators should become more familiar with body language and the meanings behind movements and facial expressions. All this awareness becomes part of the evaluation evidence.

Teachers need to look for direct evidence of children learning and be able to describe and document such occurrences. We need to rely on the informed judgment of teachers regarding the development of young children, but they (the teachers) need to become more skillful at justifying their judgments so that others (parents, administrators, etc.) will develop more confidence in that kind of subjective effort. The most appropriate means of justification is to point to concrete examples of learning that characterize the judgment rendered--examples of work, tasks performed, activities engaged in, verbal expressions by children, etc.

In developing confidence in this subjective form of evaluating children, we may be able to mitigate the impact of test data. We need to realize, and help others realize, that tests are only indirect measures of educational gains, correlates of learning rather than direct evidence of learning. It is in the children's own work and expressions that real meaningful evaluation can occur.

Becoming Accountable Through Evaluation

In a broad sense, all the activity which occurs daily in early childhood education programs needs continuous assessment. Learning and teaching rest quite heavily on the notion of evaluation. There are many factions in the program and extended program population that need to be accountable for evaluative functions. In this network of accountable forces each group needs to be open and honest in formulating their judgments. The individual child in an early childhood program should learn to define learning objectives for himself and develop criteria with which to evaluate his progress. He must continually synthesize his new discoveries or failures into a perceptual understanding of the world around him. Programs for young children need to help each child maximize his motivational state so that he can fulfill his own purposes, goals, and needs. Every child should be encouraged to recognize the congruency between what he perceives to do and what he is actually engaged in. Early childhood educators should be accountable for involving children in this kind of evaluation process.

Within this accountability scheme the early childhood teacher has great responsibility. It is the teacher who needs to develop an understanding of each child to be able to relate on a specific basis. In addition, she should be knowledgeable and versed in both the process of learning and the developmental structure of a child's growth. This should be understood fully yet open to re-evaluation and strategic consideration. The teacher should encourage and assist each child toward maximizing his or her own potential. And it is in this process that the diagnostic activity occurs.

But the teacher is not the only one who is accountable within this system. The administrator, paraprofessionals, parents, and community groups are all accountable for making a child's experience in an early childhood program a truly rewarding one. Those groups and individuals must allow the program to become an institution of change: a program that is truly responsive to the individual child. They must all share the responsibility of allowing the evaluation procedures to respect the integrity of children. They must all be patient but, above all, they must be flexible to maintain a program which is not only responsive to innovations in early childhood education, but which can actually cause those innovations to occur.